DOCUMENT RESUME

ED 457 219 TM 033 317

AUTHOR Miller-Whitehead, Marie; Abbott, Gypsy
TITLE Efficacy, Outcomes, and Empowerment Evaluation of a School District NET Project, Part III: Theoretical Model.
INSTITUTION Tennessee Valley Educators for Excellence, Muscle Shoals, AL.
PUB DATE 2001-08-17
NOTE 21p.; For Part I of this study, see TM 033 189; for Part II, see TM 033 257.
AVAILABLE FROM Tennessee Valley Educators for Excellence, P.O. Box 2882, Muscle Shoals, AL 35662 ($5). Web site: http://www.tvee.org.
PUB TYPE Reports - Research (143)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Attitude Change; Community Attitudes; Community Change; Curriculum; *Health Education; *High School Students; High Schools; Knowledge Level; Models; *Nutrition Instruction; Outcomes of Education; Program Effectiveness; Program Evaluation; *Research Design; Student Attitudes; Student Empowerment; Theories

ABSTRACT

The efficacy and outcomes of a U.S. Department of Agriculture Nutrition Education Training (NET) project designed to increase awareness of health risk and wellness factors for grade 9 students were studied. This paper comments on results from a survey of 125 ninth graders about the NET curriculum and results from a study of program impact on schools and the community. In this section of the report, the focus is on the theoretical models and research approaches underlying the study. For this NET initiative, there were several project design and implementation concerns to be addressed that were generic to program evaluation or specific to the NET program directives. For the NET project a single group time series design was determined to be an acceptable way to examine changes in student nutrition behaviors. The NET program directives required the use of a needs assessment student survey, and models for handling missing survey data were selected from software packages commonly available. Facet theory was used to quantify complex causal paths through mapping sequences developed to describe facets for involvement of school and community resources. Cost effectiveness was another factor to be considered in study design. The results of this NET project indicated significant improvement in reported nutrition behaviors for the students at the two schools with the most students at risk for poor nutrition. Positive results at these schools were an indication of cost effectiveness for the program as a whole. (Contains 3 figures, 3 tables, and 35 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 457 219

Efficacy, Outcomes, and Empowerment Evaluation of a School District NET Project Part III:

Theoretical Model

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Marie Miller-Whitehead

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Marie Miller-Whitehead
TVEE.ORG

Gypsy Abbott
University of Alabama at Birmingham

August 17, 2001

TM033317

Direct correspondence to:

Marie Miller-Whitehead, Ph.D.
Executive Director
Tennessee Valley Educators for Excellence
PO Box 2882
Muscle Shoals, AL 35662

BEST COPY AVAILABLE

Efficacy, Outcomes, and Empowerment Evaluation of a School District NET Project Part III: Theoretical Model

Introduction

Programs designed to increase health awareness in adolescents and to effect positive changes in adolescent health behaviors may pose special challenges to the evaluator, particularly if the activities are school-based as was the case with a district-wide grade 9 Wellness NET project. NET (Nutrition Education Training) is supported by the USDA and is a component of the state and federal school lunch program. Fortunately, a project director has several options that will not only satisfy the requirements of the program's funding agency but will also provide an evaluation helpful to the stakeholders and to the population served. The most widely accepted standards for program evaluation address four areas: utility, feasibility, propriety, and accuracy (Joint Committee on Standards for Educational Evaluation, 1994). These standards provide a broad framework for program evaluations and will be referred to later.

The development of heuristics specific to a particular category of program has received attention from many evaluation methodologists: an evaluation model developed primarily for wellness and nutrition programs that target adult populations may not be suitable within the context of the K12 environment, and within the K12 framework not all wellness and nutrition program evaluation models are equally proper or feasible due to a variety of constraints that may be operational within either the school, the community, or the requirements of the program to be evaluated. In this context the evaluator must differentiate between models appropriate for the evaluation of social programs and those more appropriate for the evaluation of student learning (Greene, 2000; Guba & Lincoln, 1981; Joint Committee, 1994). As was the case with the NET grade 9 project, most state and national programs are evaluated by an external evaluator and require or suggest the use of a specific evaluation model and measurement instruments. The NET program was goal-oriented and had as its focus a limited set of target objectives that reflected the Dietary Guidelines for Americans with participants expected to improve in attainment of healthy food consumption habits. Goal-oriented evaluation measures progress to an identified set of criteria (Popham, 1988). Program specified objectives and reporting procedures assure consistency in reporting of results across the various populations that will be served by funded programs, may simplify or eliminate the necessity of obtaining informed consent by participants and eliminate a certain amount of other red tape , and can be cost effective in allowing more funds to be directed toward achieving results for program participants at the local level (Ceci, Peters, & Plotkin, 1985; Kraemer, 1981; Nas, 1996; Yates, 1996). However, evaluation models provided by funding agencies may be frameworks that are designed more to serve the accountability requirements and goals of the parent organization than to provide the most useful feedback to program participants and community stakeholders. Reporting of results for the NET program appeared to be designed to facilitate synthesis and meta-analysis of district level projects by the external state and federal evaluation team (Glass, McGaw, & Smith, 1981). However, participants in community and school district projects will often not only wish to know "how did we do?" at the end of a project but they may also wish to

know "how did we do compared to other schools and school districts in our state or across the nation, or to our own past performance on these objectives?" Although neither participant self-evaluation nor stakeholder involvement in production of an evaluation product reflective of community participant goals and objectives is a stand-alone substitute for a professional external evaluation, together they form the guiding principle of empowerment evaluation or goal-free evaluation, approaches to evaluation that have been embraced by many community agencies (Fetterman, Kaftarian, & Wandersman, 1996; Scriven, 1991). If a project is to be continuous rather than a one-shot effort, lack of evaluation information that is appropriate to the goals and objectives of the local community stakeholders may lead to difficulties in future recruitment of teachers and community organization participants at the local level (Harrington et al., 1997). Scriven's goal-free evaluation is not constrained to reporting only whether a project achieved a set of a priori objectives, but rather focuses on what actually occurred as a result of the program (Maduas, Scriven, & Stufflebeam, 1983; Scriven, 1991; Shadish, Cook, & Leviton, 1991). For example, research has shown that repeated interventions and training did not improve the nutrition behaviors of female athletes, even when their knowledge of nutrition increased significantly as a result of their training (Chapman, Toma, Tuveson, & Jacob, 1997; Edwards, Acock, & Johnston, 1985). Other studies have shown that better educated people with higher incomes scored higher on the USDA's Healthy Eating index than did less well-educated people with lower incomes, with African-Americans generally having lower scores on the Healthy Eating index than other ethnic groups and Southern rural populations generally having lower scores than other regions of the nation (Bowman, Lino, Gerrier, & Basiotis, 1998). Without adequate comparison data and incorporation of some elements of goal-free or empowerment evaluation, community participants might fail to credit any level of success to their project if results showed only slight or insignificant improvement in target objectives. Stufflebeam's (1994) recommendations for improving evaluation in schools call for the use of both prospective and retrospective context evaluation to determine the worth of a program in meeting the assessed needs within the context of the student population served.

For the grade 9 NET school district initiative there were several project design and implementation concerns either generic to program evaluation in general or specific to the NET program directives. These involved questions related to selection of student population and number of participants, necessity (or not) of obtaining parental permission or informed consent, development of a realistic project timeline, identification of achievable project objectives, assurance of adequate and appropriate school district and community involvement, assurance of adequate documentation and compliance with NET objectives, selection of best methodology for measurement of project objectives within the framework of the project timeline, and measurement and communication of project's benefits and cost effectiveness to the community and to the participants (Guba & Lincoln, 1981; Joint Committee, 1994; King, Morris, & Fitz-Gibbon, 1987; Nas, 1996; Patton, 1987; Stufflebeam, 1994; Yates, 1996). Table 1 (Greene, 2000, p. 984) provides descriptions of the most widely-used contemporary evaluation models, with the state requirements of the NET program incorporating some elements of utilitarianism but being primarily reflective of the postpositive or goal-oriented model.

Insert Table 1 about here

There were several choices available in selection of the target population in that the project could have been designed to serve Title 1 students and their families, clients of the Family Resource Center, or grade 9 Wellness students: there were demonstrable needs among each of these populations. Project population selection was based on the results of a preliminary needs assessment conducted among students in two school systems, the fact that Tennessee had recently adopted a new grade 9 Wellness curriculum framework, the fact that the teachers and school libraries had limited ancillary materials aligned to the new curriculum, and the knowledge that a project designed to support the grade 9 Wellness program would serve over 1,500 students (some of whom were also Title 1 students) while there were less than 100 students and families who would have been served had the project targeted Title 1 or the community Family Resource Center. Funding agencies often consider not only results of a needs assessment but also numbers of participants to be served and thus the decision was made to design the program for the grade 9 Wellness students with provision for materials to be made available to the Family Resource Center. In the presence of other constraints (such as problems with obtaining permissions, random assignment of participants, and maintaining confidentiality in working with the smaller group of Title 1 or Family Resource Center students) the selection of the larger population is the usual choice when the researcher has empirical evidence that an intervention is likely to have a small effect (Rosnow & Rosenthal, 1988).

Insert Figure 1 about here

For the NET project a single group time series design was determined to be an acceptable way to examine changes in grade 9 student nutrition behaviors for several reasons, not least of which related to obtaining permissions and likelihood of contamination for a design using a treatment and control group approach within Wellness classes at five high schools in the same school district (Ceci et al., 1985; Fitz-Gibbon & Morris, 1996). According to Cook and Campbell (1979), a time series analysis (Figure 1) has several major advantages over other quasi-experimental designs, allowing assessment of maturational trends before and after a course of treatment. Mohr (1995) observed several possible sources of error in measurement of the difference in two points, or the before-and-after design, compared to the use of the interrupted time series models, due primarily to the inability of the former to account for a trend or maturation effect, or the possibility of an abrupt change in intercept or slope (Figure 2). In the case of the NET program, this could be the result of menu changes in the school cafeteria, snack selections in the school vending machines, seasonable availability of foods, or presence or absence of school fundraising activities such as candy sales.

The design plan called for each Wellness class to have at least two and preferably three nutrition interventions (speakers, projects, cooking labs, or special activities such as field trips) between pre- and posttesting with the assumption that additional interventions would reinforce

knowledge and positive behaviors prior to the posttest. Herman, Morris, and Fitz-Gibbon refer to this as the before and after design (p. 113). The literature suggested the existence of trends in adolescent nutrition habits over time, with younger children having healthier nutrition habits than teens; thus comparison of the NET project trend data with state and national indicators in teen nutrition was of interest (Bowman et al., 1998; Health-risk behaviors, 1994). Available data on comparable measures from the Youth Risk Behavior Survey and from the National Longitudinal Study of Adolescent Health also indicated an increase in some risk factors for adolescents in the early teen years. Nutrition education is a strand of the grade 9 Wellness curriculum in Tennessee and is a required course for all grade 9 students. While the program design was not experimental, random assignment was more likely to occur in grade 9 Wellness classes than in many other classes because the school district's Wellness courses were not ability tracked at any of the five high schools in this study and all grade 9 students regardless of ability were required to enroll in a Wellness class. Thus, Wellness classes included all but the most severely disabled students, representing a random sample of the population of grade 9 students in each school within the district. With 125 pretests and 100 available posttests, the total surveyed population represented a matrix sample of 225 grade 9 students for whom responses were processed.

Insert Figure 2 about here

The NET program directives required use of a needs assessment student survey. The use of survey instruments with student populations is often restricted not only by cost and considerations for confidentiality and privacy, but also by sometimes conflicting district, state, and federal standards, particularly when students are asked to self-report health risk behaviors such as alcohol, tobacco, or drug use (Ceci et al., 1985; Henerson, Morris, & Fitz-Gibbon, 1987; Jaeger, 1997; Joint Committee, 1994). While not an ideal solution, based on considerations of political viability and potential questions about the rights of human subjects, the decision was made to pre- and posttest a random sample of grade 9 Wellness students with responses only coded by gender, age, and school assuring students of complete individual anonymity with no coding or identification of individuals. While this precluded the analysis of paired samples by individual, pre- and posttest comparisons by gender and by school were conducted, with school a proxy for community characteristics such as SES and education level. Project sites were five high schools in a large county with significant differences in average student SES by school location. In this case, in addition to confounding issues related to confidentiality, there were issues related to the sampling of the student population. Enrollment of grade 9 students at the five high schools ranged from 185 to 541 students, with a total of between 1400 and 1500 grade 9 students who would receive additional nutrition education training and support from the NET program. In this case, two classes or approximately 50 students were surveyed from each of the five high schools, for a total of between 225 and 250 student surveys for both pretest and posttest analyses. Although small sample research is often a reality, 30 members per group, 20% of the population, or 200 members of a population are often recommended to assure sampling adequacy (Kraemer, 1981; Slavin, 1992). Due to constraints of time and budget the pre- and posttest analyses were conducted with a random sample of 25 surveys from each school (100 total), with a sample size of 20 per group being deemed adequate for this type of assessment

(Kraemer; Slavin). As pointed out by Slavin, "if the study will cost the district any money, its chances of being approved are minimal" (p. 126). While a larger sample size might have yielded larger effect sizes for pre- and posttest differences or have been more powerful to detect differences, it would have necessitated the hiring of additional personnel to handle clerical duties, which would likely not have met with the approval of the district's funding agencies. In these cases the evaluator must determine the relative expense and risk involved in making a Type I error (probability of rejecting the null hypothesis when it is true) or Type II error (probability of rejecting the null hypothesis when it is false). In a two-tailed test at an alpha level of .05 for the t test of independent samples, the power of a study conducted with 20 subjects per group for an effect size of .5 is a rather low .33, with 50 subjects per group and an effect size of .5 power increases to .70, and with 100 subjects per group, power increases to .94 (Stevens, 1990). Thus, the design of the NET project was powerful for detecting pre- and posttest between group differences and within group gender differences, although there was a only 33% chance of detecting differences that may have existed between the five high schools of this study. With 20 to 25 subjects per group, conducting a unidirectional test would have increased power substantially.

Although there are a variety of ways to address the issue of missing data and most statistical software packages offer a choice of replacing the missing values with the either the mean or deletion, surveys with missing data are routinely deleted from analyses under the assumption that missing data are randomly distributed across populations of interest and thus their deletion will not effect the results of the analysis. However, the researcher will often wish to conduct a separate analysis of surveys or tests with missing data as the results may provide information of interest to the researcher. Of the 41 item Wellness survey, several items asked students to respond to various questions about personal values and health risk behaviors such as alcohol, tobacco, or drug use. Although responses were completely anonymous, students were instructed to leave any item blank if they felt their parents would object to the question. Thus, surveys with missing responses for certain questions would be assumed to represent a population of students not representative of the normal distribution for those items. Surveys with missing values were subsequently analyzed separately to identify relationships between various patterns of nonresponse and to determine if the use of an imputation model would be effective. Many of the latest versions of statistical software packages provide several models for handling missing data using various regression procedures to estimate missing values. There were 13 surveys with missing data for the gender variable; correlation coefficients and regression models identified a combination of several variables that were correlated with the gender of the respondent for the surveys with the gender variable coded. These variables were subsequently entered as covariates in a logistic regression model which correctly identified 88% of female respondents but only 54% of respondents for whom gender data were missing, an indication supportive of a hypothesis that data missing the code for gender were randomly distributed.

Insert Table 2 about here

The Logic Model

Cook (2000) points out three steps common to most theory-based logic models: first, the development of the theoretical model and anticipated participant achievement, second, the collection of data sufficient to measure outcomes, and third, the analysis of the data to determine the extent to which anticipated outcomes occurred within the predicted time frame of the program. The inherent difficulty with the theory-based logic model is that it presumes continuity in the causal chain of events and assumes to a great extent that programs occur in a vacuum when in reality there are likely numerous "uncontrolled local details that provide contextual variables interacting with treatments" (Cronbach, p.329). Cronbach advises evaluators to spend less time worrying about the null hypothesis and more time identifying and controlling confounding variables. Not uncommonly project outcomes can best be determined in the context of a retrospective evaluation (Stufflebeam, 1994). For example, compare the postulated simple pre- and posttest model of Figure 1 (a leading period model to predict participant outcome based on two interventions) with the retrospective model of Figure 3. Figure 3 provides an example of a project in which participants hypothetically were demonstrated to have achieved significant improvements well after the end of the end of the project due to delays in the planned interventions. These results were indicated by an ARIMA model study conducted the following year and post project interviews (Cook & Campbell, 1979).

Insert Figure 3 about here

Since there are both monetary and other consequences for participants and funding agencies when projects are determined to have achieved or not achieved their objectives, an examination of causality is of interest. If a high correlation has been identified between a program intervention (A) and a dependent variable such as a positive change in nutrition behavior (B) there still must be additional empirical support for the evaluator to assert with any degree of confidence that A caused B. Determination of causality has been the object of extensive methodological research and debate and there are numerous approaches to identification of a causal chain or causal paths, including the propriety of identifying all possible side effects or unintended consequences that may be related to the program under examination (Davidson, 2000). Devout proponents of experimental design often find fault with structural equation models as they do not require manipulation of the dependent variable and thus infer causality through covariance; however, structural equation models are often preferred to experimental design in clinical studies for a variety of reasons, particularly in treatment of children where it is not desirable to withhold a treatment or intervention as a control (Morris, Bergan, & Fulginiti, 1991). Davidson (p. 21) has identified the following nine types of evidence that support the inference of causality:

1. Causal list inference (almost all Bs are caused by A, A', A")
2. Modus operandi inference (more than one possible cause has occurred, A was present)
3. Temporal precedence (A happens before B)
4. Constant conjunction (when A, always B)

5. Contiguity of influence (there is a logical link between A and B)
6. Strength of association (much more B with A than with other possible causes)
7. Biological gradient (if more A, then more B)
8. Coherence (the A-B relationship fits with everything else we know about A and B)
9. Analogy (the relationship of A and B resembles the established pattern of C and D)

Facet theory, based on the work of Guttman, provides another approach used in social science contexts to quantify complex causal paths through mapping sentences (Borg & Shye, 1995). Facets for the grade 9 student participants might include initial status on the specified dimension by gender, at specified age, as a resident of a particular community or student at a specified school (the SES proxy), for cognitive, affective, or behavior level, provided x number of interventions of a specified type or combination of types. Thus, if two kinds of interventions are possible and there are two interventions provided by the program, the student might be provided one type at time 1 and the second at time 2, hoping for a multiplier effect or interaction in which the combination of the interventions would be more effective than each intervention measured individually. A mapping sentence for the NET program objectives might contain the following facets for student participants:

R1 = (good) ( ) (poor) Health risk behavior

The level of (c1 = cognitive) (c2 = affective) (c3 = behavioral) functioning at age (13) (14) (15) at school (A) (B) (C) (D) (E)

for (g1 = male) (g2 = female) on

Dimension
(nutrition)
(physical activity)
(substance abuse)
(traffic safety)
(personal values)
(self-care)

with intervention

Kind of Intervention
(b1 = ecological)
(b2 = psychological)

at time (t) → = R2

Because the NET program directives mandate community involvement the mapping sentences were developed to describe facets for involvement of school and community resources. Aspects included group membership, perceived difficulty in effecting change on a specified nutrition dimension for a particular type of intervention, for male or female students who attended a specified school or who were members of a certain community, provided with x number of interventions. The resultant proposed sentence mapping for perception of school and community participants is provided below:

Person (p), who is a

Membership		Positively		Dimension
(m1 = teacher)	believes it is	(p1 = easy)	to change	(Nutrition)
(m2 = care provider)		(p2 = neutral)		(Physical Act)
(m3 = dietician)		(p3 = difficult)		(Substance Abuse)
(m4 = cafeteria personnel)				(Traffic Safety)
(m5 = social service professional)				(Personal Values)
(m6 = school administrator)				(Self-Care)

	Type of Intervention				Gender
at level	(i1 = cognitive)	for adolescents at age	(13)	who are	(g1 = male)
	(i2 = affective)		(14)		(g2 = female)
	(i3 = behavioral)		(15)		
			(16)		

	Location	
students at school	(A)	at time (t)
	(B)	
	(C)	
	(D)	
	(E)	

	Type of Resource
given resource	(b1 = ecological)
	(b2 = psychological)

Results and Cost Benefit and Cost Effectiveness

Programs and evaluations should be cost effective. Local supporters and participants do not want to feel that they have wasted time or other resources better spent in other activities or projects. As reported by Harrington et al., while most teachers are very responsive to participation in projects that will potentially benefit their students, care must be taken in intruding on their class time to survey or assess students. Additionally, schools and school districts do not wish to feel that agreeing to participate in a research project has cost them any

money not reimbursed by project funds (Slavin, 1992). The determination of numbers of participants necessary is dependent not only upon purpose of the study but also upon the expense of recruiting and processing participants and whether it is advisable to minimize Type I or Type II error (Kraemer, 1981). Funding agencies that award grants cannot fund all applicants and projects that receive continuing funding must be prepared to show that their project provided an adequate return on investment for the funding agency such that the project justified its existence. A goal-oriented project should be able to show either achievement of the goal or progress to the goal comparable to or exceeding that of nonparticipants. There is a difference between cost benefit and cost effective measures. Cost benefit analysis involves assignment of a monetary value to the project outcomes while cost effectiveness analysis refers to valuation of outcomes that cannot be easily assigned a monetary value. Since relatively few grade 9 students incur monetary loss or expense due to poor nutrition habits it is difficult to place a monetary value on project outcomes. When students do have the unfortunate experience of having a classmate become very ill or perhaps die due to poor nutrition choices or eating disorders it may serve as an incentive to both the students and the community to support efforts in raising awareness of the importance of good nutrition for adolescents. Long-term health benefits for students and potential monetary benefits in reduced future health care costs far outweigh project expenses if there is empirical evidence of improved nutrition habits for student participants, and if even one teen can be spared a catastrophic illness the project will have had a positive benefit to cost ratio.

Although male students showed more improvement than female students, the results of this NET project indicated significant improvement (Table 3) in reported nutrition behaviors in grade 9 students at the two high schools with most students at risk for poor nutrition either because of SES or ethnicity (Bowman, Lino, Gerrier, & Basiotis, 1998). Since both male and female African-American adults experience coronary heart disease and adult diabetes both more frequently and at a younger age than other groups, the positive results of the NET project for grade 9 students at schools with larger percentages of African-American students provided an indication of cost-effectiveness for the program, with a variety of resource materials targeted to these groups placed in the school libraries and at the Family Resource Center, which served not only a large African-American population but also a large population of ESL students and their families.

Bibliography

Borg, I., & Shye, S. (1995). Facet theory: Form and content. Advanced Quantitative Techniques in the Social Sciences Series. Thousand Oaks: Sage Publications.

Bowman, S. A., Lino, M., Gerrier, S. A., & Basiotis, P. P. (1998). The Healthy Eating Index 1994-96. U.S. Department of Agriculture, Center for Nutrition Policy and Promotion. CNPP-5.

Ceci, S. J., Peters, D., & Plotkin, J. (1985). Human subjects review, personal values, and the regulation of social science research. Reprinted in A. E. Kazdin, (Ed.). (1992), Methodological issues and strategies in clinical research (pp. 687-704). Washington, D.C.: American Psychological Association.

Chapman, P., Toma, R. B., Tuveson, R. V., & Jacob, M. (1997). Nutrition knowledge among adolescent high school female athletes. Adolescence, 32 (126), 437-446.

Cook, T. D. (2000). The false choice between theory-based evaluation and experimentation. New Directions for Evaluation, 87, 27-34.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design & analysis issues for field settings. Boston: Houghton Mifflin.

Cronbach, L. J. (1991). Functional evaluation design for a world of political accommodation. In W. R. Shadish, Jr., T. D. Cook, & L. Leviton (Eds.), Foundations of program evaluation: Theories of practice, pp.323-376. Newbury Park, CA: Sage Publications.

Davidson, E. J. (2000). Ascertaining causality in theory-based evaluation. New Directions for Evaluation, 87, 17-26.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2000). Handbook of qualitative research. (2nd ed.). Thousand Oaks, CA: Sage Publications.

Edwards, P. K., Acock, A. C., & Johnston, R. L. (1985). Nutrition behavior change: Outcomes of an educational approach. Evaluation Review, 9, 441-460.

Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (Eds.). (1996). Empowerment evaluation: Knowledge and tools for self-assessment and accountability. Thousand Oaks: Sage Publications.

Fitz-Gibbon, C. T., & Morris, L. L. (1987). How to design a program evaluation. Newbury Park, CA: Sage Publications, Inc.

Glass, G. V, McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hill, CA: Sage Publications.

Greene, J. C. (2000). Understanding social programs through evaluation. In Denzin, N. K., & Lincoln, Y. S. (Eds.). (2000). Handbook of qualitative research. (2nd ed.). (pp. 981-1000). Thousand Oaks, CA: Sage Publications.

Guba, E. G., & Lincoln, Y. S. (1981). Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches. San Francisco: Jossey-Bass.

Harrington, K. F., Binkley, D., Reynolds, K. D., Duvall, R. C., Copeland, J. R., Franklin, F., & Raczynski, J. (1997). Recruitment issues in school-based research: Lessons learned from the High 5 Alabama Project. Journal of School Health, 67(10), 415-421.

Health-risk behaviors among persons aged 12-21 years-United States, 1992. (1994). Reprinted from Centers for Disease Control Morbidity and Mortality Weekly Report. In Journal of the American Medical Association, 271(17), 1311.

Henerson, M. E., Morris, L. L., & Fitz-Gibbon, C. T. (1987). How to measure attitudes. UCLA. Newbury Park, CA: Sage Publications.

Herman, J. L., Morris, L. L., & Fitz-Gibbon, C. T. (1987). Evaluator's handbook. Newbury Park, CA: Sage Publications, Inc.

Jaeger, R. M. (Ed.) (1997). Complementary methods for research in education. (2nd ed.). Washington, D.C.: American Educational Research Association.

King, J. A., Morris, L. L., & Fitz-Gibbon, C. T. (1987). How to assess program implementation. Thousand Oaks, CA: Sage Publications, Inc.

Kraemer, H. C. (1981). Coping strategies in psychiatric clinical research. Reprinted in A. E. Kazdin, (Ed.). (1992), Methodological issues and strategies in clinical research, (pp. 539-556). Washington, D.C.: American Psychological Association.

Madaus, G., Scriven, M., & Stufflebeam, D. (1983). Evaluation models. Boston, Kluwer Nijhoff.

Mohr, L. B. (1995). Impact analysis for program evaluation. Thousand Oaks, CA: Sage Publications.

Morris, R. J., Bergan, J. R., & Fulginiti, J. V. (1991). Structural equation modeling in clinical assessment research with children. Reprinted in A. E. Kazdin, (Ed.). (1992), Methodological issues and strategies in clinical research, (pp.367-386). Washington, D.C.: American Psychological Association.

Nas, T. F. (1996). Cost-benefit analysis: Theory and application. Thousand Oaks, CA: Sage Publications.

Patton, M. Q. (1987). How to use qualitative methods in evaluation. Newbury Park, CA: Sage Publications.

Popham, W. J. (1988). Educational evaluation. (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Rosnow, R. L., & Rosenthal, R. (1988). Focused tests of significance and effect size estimation in counseling therapy. Reprinted in A. E. Kazdin, (Ed.). (1992), Methodological issues and strategies in clinical research, (pp.337-349). Washington, D.C.: American Psychological Association.

Scriven, M. (1991). The science of valuing. In W. R. Shadish, Jr., T. D. Cook, & L. Leviton (Eds.), Foundations of program evaluation: Theories of practice, pp. 73-118. Newbury Park, CA: Sage Publications.

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). Foundations of program evaluation: Theories of practice. Newbury Park, CA: Sage Publications.

Slavin, R. E. (1992). Research methods in education. (2nd ed.). Needham Heights, MA: Allyn and Bacon.

Stevens, J. (1990). Intermediate statistics: A modern approach. Hilldale, NJ: Lawrence Erlbaum Associates.

Stufflebeam, D. L. (1994). Introduction: Recommendations for improving evaluations in U.S. public schools. Studies in Educational Evaluation, 20, 3-21.

Yates, B. T. (1996). Analyzing costs, procedures, processes, and outcomes in human services. Applied Social Research Methods Series. (Vol. 42). Thousand Oaks, CA: Sage Publications.

O_1	O_2	X_1	X_2	O_3

Figure 1. Pre- and posttest design with two interventions, where O_1 = needs assessment, O_2 = pretest, X_1 = psychological intervention followed within two week period by X_2 (ecological intervention) and O_3 = posttest.

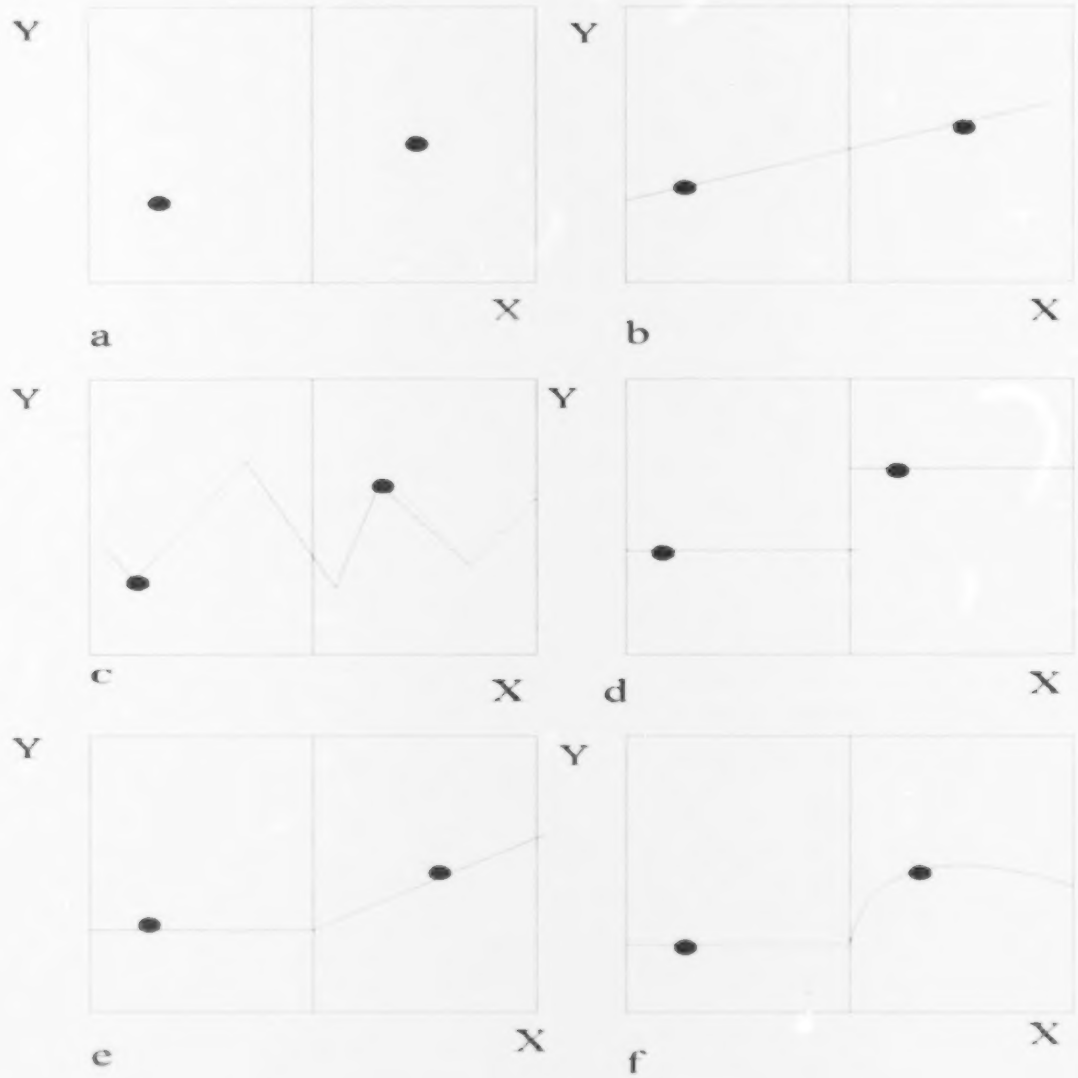

Figure 2. Time-Series Analysis Compared With Analysis of Two Time Points

Note. Where (a) is apparent effect, (b) is explanation by trend, (c) is explanation by noise, (d) is possible transfer function: abrupt change in intercept, (e) is possible transfer function: abrupt change in slope, and (f) is possible transfer function: gradual change in intercept (Mohr, 1995, p. 207).

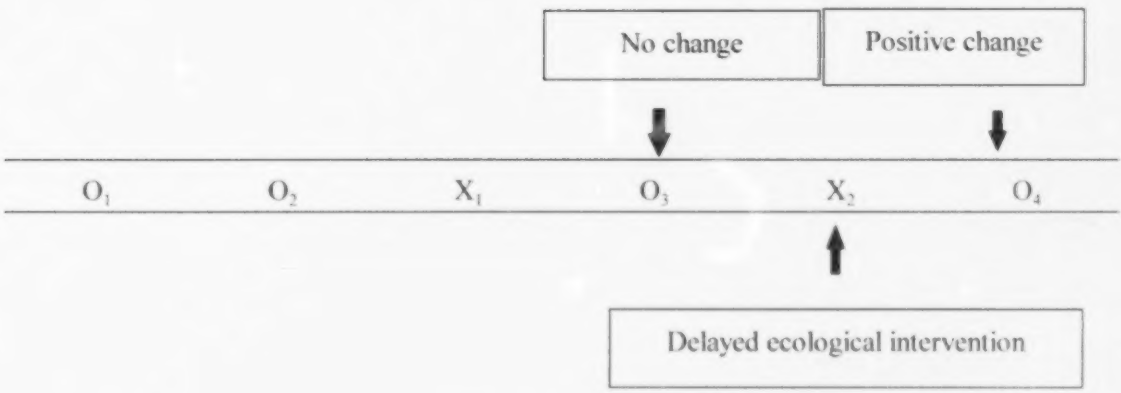

Figure 3. Retrospective time series, where O_1 = needs assessment, O_2 = pretest observation, X_1 = psychological intervention, O_3 = posttest observation, X_2 = delayed ecological intervention, and O_4 = retrospective observation demonstrating significant positive participant behavioral change

Table 1

Major Contemporary Approaches to Formal Program Evaluation[a]

Epistemology	Primary Values Promoted	Key Audien
Postpositivism (Cook, 1985)	Efficiency, accountability, cost effectiveness, policy enlightenment	High-level de fu sc co
Utilitarian pragmatism (Patton, 1997)	Utility, practicality, managerial effectiveness	Midlevel pr ma sit
Interpretivism, constructivism (Stake, 1995)	Pluralism, understanding, contextualism, personal experience	Program di an
Critical social sciences (Fay, 1987)	Emancipation, empowerment, social change, egalitarianism, critical enlightenment	Program be the act

[a] Reprinted from Greene, J. C. (2000). Understanding social programs through eva
qualitative research. (2nd ed.). Thousand Oaks, CA: Sage Publications. p. 984.

ces	Preferred Methods	Typical Questions
olicy and cision makers, ders, social ence mmunity ogram nagers and on- e administrators	Quantitative: experiments, quasi-experiments, surveys, causal modeling, CBA	Are intended outcomes attained and attributable to the program? Is this program the most cost efficient alternative?
	Eclectic, mixed: structured and unstructured surveys, interviews, observations, document analysis, panel reviews	Which program components work well and which need improvement? How effective is the program with respect to the organization's goals and mission? Who likes the program?
ectors, staff, l beneficiaries	Qualitative: case studies, open-ended interviews and observations, document reviews, dialectics	How is the program experienced by various stakeholders? In what ways is the program meaningful?
neficiaries and ir communities, ivists	Participatory, action oriented: stakeholder participation in evaluation agenda setting, data collection, interpretation, and action	In what ways are the premises, goals, or activities of the program serving to maintain power and resource inequities in this context?

luation. N. K. Denzin and Y. S. Lincoln, (Eds.). (2000). Handbook of

6

Table 2

Logistic Regression for NET Survey Variable Responses Predictive of Gender

	Predicted		
Observed	Male	Female	Percent Correct
Male	15	14	51.72%
Female	7	51	87.93%
		Overall	75.86%

Note. Test for goodness of fit, $\chi^2 = 2.6115$, df = 6, sig. = .8558

Table 3

Pre- and Posttest Means on School District FIT-KID Objectives by School

	School	Pretest (n = 25)	Posttest (n = 20)	ES
I limit high fat foods				
	A	.24	.40	.34
	B	.24	.27	.07
	C	.28	--	
	D	.48	.28	-.41
	E	.40	.32	-.16
I eat at least 5 servings of fruits, vegetables, grains daily				
	A	.12	.23	.29
	B	.16	.35	.44
	C	.24	--	
	D	.28	.56	.57
	E	.13	.21	.21

Note. The community census sector per capita income for School A and School B was lower than that for School C, School D, or School E. School A was located in a relatively rural area, while School B was within city limits. For the pretest there were 25 subjects per school and 20 subjects per school for the posttest. Posttest results were not available for School C.